Chapter 11 – Search

Terms

This Windows SharePoint Services Operations Guide (WSSOPS) from the Computer Information Agency is provided as is. Every effort has been made by the author to ensure that information provided is correct however this cannot be guaranteed.

By using the Guide you also acknowledge that any work performed on production systems is a potentially dangerous act and may involve significant business risk and downtime. You hereby agree to release, waive and discharge the Computer Information Agency and/or the author from any liability incurred to yourself, your business or customers for any and all loss or damage, and any claims or demands therefore on account of problems arising from the use of this Guide.

By using this guide you hereby assume full responsibility for any risk to computer systems, now and forever arising out of, or related to the use of this Guide on any computer system.

Distribution and Duplication Guidelines

This document is copyright and only available directly from the Computer Information Agency. This work is now licensed under Creative Commons.

http://creativecommons.org/licenses/by-nc-sa/4.0/

By using this Guide you agree to these terms.
Chapter 11 – Search

Index

11.1 Scope
11.2 Prerequisites
11.3 Basics
11.4 Ensuring the search is enabled
11.5 Indexing Acrobat PDF’s
11.6 Indexing Office 2007 documents
11.7 Manual Full Crawl
11.8 Microsoft Search Sever Express 2008
11.9 Indexing .MSG files
11.10 Conclusion
11.1 Scope

This chapter will cover the specifics of enabling and enhancing the searching of Windows SharePoint Service v3 (WSS v3). Although WSS v3 can search much of what appears in WSS v3 it is possible to extend this to items such as Adobe PDF documents. The major restriction that searching has in WSS v3 is that it is limited to the WSS v3 site. However, with the installation of Microsoft Search Server Express 2008 it is possible to not only search WSS v3 content but also information from elsewhere including network file shares, Exchange public folder and even Internet web sites.

11.2 Prerequisites

WSS v3 search requires WSS v3 to be installed. See Chapters 3 and 4 for more information about installing WSS v3.

11.3 Basics

Unlike its predecessor SharePoint v2, WSS v3 includes the ability to index and search its content automatically.

Once WSS v3 has been installed you will see a search box in the top right of the window. Simply enter the desired search information and click the magnifying glass to the right to commence the search.
When the search has completed you will a list of entries matching your search display like shown above. You can click on these results to display the entry. If you enter more than one word WSS v3 will match any item with either or both of the strings (a Boolean OR operation). WSS v3 search uses a feature called stemming. The basics of search stemming are:

- When you submit a query in SharePoint, the query is broken into individual words. For example, the query "securing the database" would be broken down into "securing", "the", and "database"
- Noise words can be eliminated, i.e. common words such as "and", "the", "or", that are unlikely to influence results. In this example, "the" would be dropped from the query. It also ignores single digit numbers.
- The query words can then be stemmed for variations. For example, a query for "security" could be expanded to include documents that refer to "securing", "securely" and so on
- The query words will also be compared against the thesaurus. The thesaurus is customizable and very useful for words with domain-specific alternatives or abbreviations. You can choose to expand queries (e.g. expand "PMB" to also search for "Purple Medium Board") or replace queries (e.g. replace "ie" with "Instant Everywhere" - "ie" will return just about every document in an English-language index).
- Does not support Boolean functions such as AND and OR. You can however include key words by using the + (plus) or – (minus) sign in front of the keyword.

The plus (+) tells the search that the word must be included in the results, and
the minus (-) indicates that the term should not be included in the results.

- It is not case sensitive.

There are however a number of components of WSS v3 that are not searchable. These include:

- Non-text columns in lists – i.e. lookup, current and Yes/No fields.
- Attachments to list items.
- Survey lists.
- Hidden lists.

By default WSS v3 will run an incremental indexing process every 5 minutes.

11.4 Ensuring the search is enabled

Depending on how you chose to install WSS v3 the search functionality may not yet be enabled.

Open the SharePoint 3.0 Central Administration via Start | Administrative Tools | SharePoint 3.0 Central Administration.
Click on the Operations tab.

Click on the Services on server link located under the Topology and Services heading.
If search isn’t running you will see a window like that shown above. Here you see the Windows SharePoint Services Search has a Status of Stopped. To enable it click on the Start link located in the Action column.

You will now be asked to enter a Service Account and a Content Access Account. Note that neither of these needs to be a domain administrator account but they must have
suitable rights to perform the requirements of the tasks listed. You may also need to use \<domain>\<username>\n
At the bottom of this screen you will find the options to configure the indexing schedule. Make any changes and press the OK button to continue.
WSS v3 will now configure the search service based on the settings you previously selected.

You now see Windows SharePoint Services Help Search is Started. You can return to this screen at any time to make changes to the search operations of your WSS v3 site.
Select the Application Management tab. Then select **Content databases** from the SharePoint Web Application Management section.

Click on the content database (in this case **sts_vmsbs2003p_1**).
Scroll down the screen until you locate the *Search Server* section. From the drop down menu select the name of your WSS v3 server.

Press **OK** to save the new configuration.

If you now return to your WSS v3 site and enter in a search request you should see the results displayed. Remember that search results will not display until an index crawl has completed.
11.5 Indexing Acrobat PDF’s

The first step in the process is to download and install Adobe PDF IFilter 6.0 from:


If there is already a version of Acrobat Reader on the server that is greater than 7.0.5 then the ifilter capability is already present. It is recommended that you upgrade any
version of Acrobat on the server to the latest version to ensure you have the latest facility for PDF document searches. If do not have Acrobat installed on your WSS v3 server then you will need to proceed with the installation of the ifilter file.

**Double click** on the downloaded file to commence the installation process.

Press **Next** to continue.
Press **Accept** to continue.

Press **OK** to complete the iFilter installation.
The next step in the process is to configure WSS v3 search to examine PDF documents. To do this you will need to run the registry editor. To do this run Regedit via Start | Run | regedit.

In RegEdit, check to see if `HKLM\SOFTWARE\Microsoft\Shared Tools\Web Server Extensions\12.0\Search\Setup\ContentIndexCommon\Filters\Extension\pdf` is present and has a GUID for a default value (in this case `{4C904448-74A9-11D0-AF6E-00C04FD8DC02}`)
Check to see if `HKLM\SOFTWARE\Microsoft\Shared Tools\Web Server Extensions\12.0\Search\Setup\Filters`.pdf is present and has:

- Default = (value not set)
- Extension = pdf
- FileTypeBucket REG_DWORD = 0x00000001 (1)
- MimeTypes = application/pdf
Open the key HKLM\SOFTWARE\Microsoft\Shared Tools\Web Server Extensions\12.0\Search\Applications\<some GUID>\Gather\Search\Extensions\ExtensionList. There should be values 0-37 (plus a default value).
With this same registry key selected, from the menu bar select **Edit | New | String Value**.

Enter the name as 38. Now double click on the new entry to edit the value.

Set this new string value to `pdf` and press **OK** to save.
Chapter 11 – Search

Exit RegEdit.

Select **Start** | **Run** | **CMD** to display a DOS prompt.

At the DOS prompt type:

```
net stop spsearch
```

followed by

```
net start spsearch
```

to stop and start the WSS v3 Search Service. Close the DOS prompt.

If you perform a search in WSS v3 any existing PDF’s in WSS v3 will not be returned, however newly added PDF’s will be once they have been indexed by WSS v3. If you modify any existing PDF in WSS v3 it will be re-indexed. To index all existing PDF content in WSS v3 you will need to complete a manual re-index of WSS v3. See the section in this chapter called *Manual Full Crawl* for details.

To add the PDF icon to WSS v3 documents firstly download the image from:

After reading the guidelines about using the icon from Adobe found here:

http://www.adobe.com/misc/linking.html#producticons

Copy the GIF file into this directory `c:\program files\common files\Microsoft shared\web server extensions\12\template\images`. 
Locate the file \c:\program files\common files\Microsoft shared\web server extensions\12\template\xml\docicon.xml. Right mouse click on the file to edit.

Locate the <ByExtension> element where you will see entries for each icon starting with <Mapping key=>. Enter the following on a new line:

```xml
<Mapping key="pdf" Value="pdficon_small.gif" OpenControl="" />
```

Note, that the extension names do not have to be in alphabetical order so it is best to place the entry at the end of the existing list. Also ensure that the correct filename for the icon is used in the entry line.

When complete, save the file and exit notepad.

To see the icon immediately in WSS v3 start a command prompt on the server via Start | Run | Cmd.
At the prompt enter `iisreset` to restart Internet Information server. When the process is complete, exit the DOS prompt and examine any WSS v3 libraries that contain PDF documents.
You should now see the PDF icon denoting the file type as PDF in WSS v3 as shown above. Please note that you are only able to open PDF documents if you have Acrobat Reader installed on the workstation.

You should now see that when you do a search your PDF document contents will be indexed as shown above.

### 11.6 Indexing Office 2007 documents

If Office 2007 documents are failing to be indexed you may need to install the Office 2007 system iFilter pack.

Firstly download the Office System 2007 filter pack from:

Press Next to continue.

Click to accept license agreement and then press the Next button.
Once this is installed according to the Microsoft deployment guide for Windows SharePoint Services found at:

http://support.microsoft.com/?id=946338

you need to complete the following steps:
You must verify and install some registry keys. To do this, follow these steps:

1. Click **Start**, click **Run**, type **regedit**, and then click **OK**.

2. Locate and then click the following registry sub-key:  
   \`\`HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Shared Tools\Web Server Extensions\12.0\Search\Applications\GUID\Gather\Search\Extensions\Extension List`\`  

3. In the details pane, verify that a key entry exists for the file name extension that you want to be indexed.

The following list contains the file name extensions that are supported by Microsoft Filter Pack:

- .Docx
- .Docm
- .Pptx
- .Pptm
- .Xlsx
- .Xlsm
- .Xlsb
- .Zip
- .One
- .Vdx
- .Vsd
- .Vss
- .Vst
- .Vdx
- .Vsx
- .Vtx

If any of these extensions do not exist, follow these steps:

a. On the **Edit** menu, point to **New**, and then click **String Value**.

b. Type the next highest supported number, and then press **ENTER**.  
   For example, if the previous highest supported number is 37, you can type **38**.

c. Right-click the registry entry that you created, and then click **Modify**.

d. In the **Value data** box, type the file format that you want, and then click **OK**.

e. Repeat these steps for any additional file name extensions that are missing.
4. In Registry Editor, locate the following registry sub key:

```
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Shared Tools\Web Server Extensions\12.0\Search\Setup\ContentIndexCommon\Filters\Extension\`
```

5. On the **Edit** menu, point to **New**, and then click **Key**.

6. Type `.docm`, and then press ENTER.
7. Right-click the registry entry that you created, and then click **Modify**.

8. In the **Value data** box, copy and paste the following information:  
   \{5A98B233-3C59-4B31-944C-0E560D85E6C3\}

9. Click **OK**.
10. Repeat steps 5 through 8 to add the following subkeys:

<table>
<thead>
<tr>
<th>Key</th>
<th>Value data</th>
</tr>
</thead>
<tbody>
<tr>
<td>.Docx</td>
<td>{5A98B233-3C59-4B31-944C-0E560D85E6C3}</td>
</tr>
<tr>
<td>.Pptm</td>
<td>{DDFE337F-4987-4EC8-BDE3-133FA63D5D85}</td>
</tr>
<tr>
<td>.Pptx</td>
<td>{DDFE337F-4987-4EC8-BDE3-133FA63D5D85}</td>
</tr>
<tr>
<td>.Xlsm</td>
<td>{F90DFE0C-CBDF-41FF-8598-EDD8F222A2C8}</td>
</tr>
<tr>
<td>.Xlsx</td>
<td>{F90DFE0C-CBDF-41FF-8598-EDD8F222A2C8}</td>
</tr>
<tr>
<td>.Xlsb</td>
<td>{312AB530-ECC9-496E-AE0E-C9E6C5392499}</td>
</tr>
<tr>
<td>.Zip</td>
<td>{20E823C2-62F3-4638-96BD-90F4F6784EBC}</td>
</tr>
<tr>
<td>.One</td>
<td>{B8D12492-CE0F-40AD-83EA-099A03493F1}</td>
</tr>
<tr>
<td>.Vsd</td>
<td>{FAEA5B46-761B-400E-B53E-E805A97A543E}</td>
</tr>
<tr>
<td>.Vss</td>
<td>{FAEA5B46-761B-400E-B53E-E805A97A543E}</td>
</tr>
<tr>
<td>.Vst</td>
<td>{FAEA5B46-761B-400E-B53E-E805A97A543E}</td>
</tr>
<tr>
<td>.Vdx</td>
<td>{FAEA5B46-761B-400E-B53E-E805A97A543E}</td>
</tr>
<tr>
<td>.Vsx</td>
<td>{FAEA5B46-761B-400E-B53E-E805A97A543E}</td>
</tr>
<tr>
<td>.Vtx</td>
<td>{FAEA5B46-761B-400E-B53E-E805A97A543E}</td>
</tr>
</tbody>
</table>

**Important** If you restart the computer before you follow these steps, the registry changes will not persist.

Exit RegEdit.

Select **Start | Run | CMD** to display a DOS prompt.
Chapter 11 – Search

At the DOS prompt type:

    net stop spsearch

followed by

    net start spsearch

to stop and start the WSS v3 Search Service. Close the DOS prompt.

11.7 Manual Full Crawl

To launch a full manual search crawl of the entire WSS v3 content complete the following steps:

1. Run a command prompt via **Start | Run | Cmd**
2. At Dos prompt type:

    C:

Then press the **ENTER** key. Then type:

    cd\Program Files\Common Files\Microsoft Shared\web server extensions\12\BIN

and again press the **ENTER** key.
Now type:

```
stsadm –o spsearch –action fullcrawlstop
```

and press the ENTER key

now type:

```
stsadm –o spsearch –action fullcrawlstart
```

and press the ENTER key

The complete re-index of all the content may take a while to complete. You will also notice an additional load on your server during the process.

### 11.8 Microsoft Search Server Express 2008

The ability to search WSS v3 content is provided by default, however it is limited to only the content contained within WSS v3 sites. With the installation of Windows Search Server Express 2008 it is possible to not only search WSS v3 content but also network file shares, web sites and Exchange Public folders.

One down side with Search Server Express 2008 is that it doesn’t integrate into the existing search functionality in WSS v3. This means that if you use WSS v3 search you
will only be able to search WSS v3 content, however if you use Search Server Express 2008 you will be able to search WSS v3 content and network files shares, web sites and Exchange Public Folders but it will be via the Search Server Express 2008 web site not the WSS v3 web site.

Microsoft Search Server 2008 Express is a freely available from Microsoft and can be downloaded from:


The specifications for Search Server Express 2008 are as follows:

- Processor: 2.5 gigahertz (GHz) (minimum)
- Memory: 2 gigabytes (GB) RAM (minimum)
- Hard disk: NTFS-formatted partition with 3 GB of free space (minimum)
- Microsoft .NET Framework 3.0 with ASP.NET 2.0 enabled
- Windows SharePoint Services 3.0
- Microsoft Windows Workflow Foundation Runtime Components

In theory it is possible to install Search Server Express 2008 onto Windows Small Business Server 2003 (SBS) but considering all the other applications that need to reside on an SBS server (including WSS v3) there is little doubt that performance of Search Server Express 2008 and SBS will be impacted if they co-exist.

The general recommendation then is to install Search Server Express 2008 and WSS v3 on a dedicated member server in a network. For more information about installing WSS v3 see Chapter 3 – Installing on Windows Server 2003 and Chapter 4 – Installing on Windows Small Business Server.

Once WSS v3 is operational on the server you then need to ensure that Windows SharePoint Services v3 Service Pack 1 is installed. With this complete you can now commence the installation of Search Server Express 2008.
Log onto the WSS v3 server as an administrator and double click on the file to commence the installation. The required files will firstly be extracted to the server.

You should now see a Start screen displayed. From here select the option **Run the Search Server Preparation Tool**.
You should now see the welcome screen. Press **Next** to continue.

The server will now be checked to ensure that it meets all the minimum requirements and is configured correctly.
When the process is complete you may need to take further action as recommended by the tool. Press **Finish** to complete the process.

You will now be returned to the **Start** window. From here now select the option **Install Search Server**.
Check the option *I accept the terms of this agreement* and press **Continue**.
Unless Search Server Express 2008 detects a compatible version of SQL server already installed, the only option that will be available on the Server Type tab will be Stand-alone. Note in that case Search Server Express 2008 will automatically also install SQL 2005 Express onto the server in the system partition. For more information about this version of SQL see Chapter 8 – Database Operations.

If Search Server Express 2008 does detect a compatible SQL version it will automatically use that. In that case you will have the option to do a Complete or Web Front end. In that case it is recommended that you select the Complete option.

If you select the File Location tab you can nominate into which directory the application files and index files will be stored. It is good practice to store these on a non system partition.

When your configuration is complete press the Install Now button to continue.
The installation process will commence and no further input will be required.

When the installation process is complete you will be given the option by default to Run the SharePoint Products and Technologies Configuration Wizard now. To do so simply leave the option checked and press the Close button.
You will then be prompted that the following services will be started or reset during the configuration:

- Internet Information Services
- SharePoint Administration Service
- SharePoint Timer Service

This may affect other users on the network and if so you should prepare them for interruption to their normal programs. Click the Yes button to accept this warning and continue. Then press the Next button to proceed.
Depending on your WSS v3 installation, Search Server Express 2008 may detect that a web site (usually WSS v3) already exists at the default location. In this case it is recommended that you select the second option which is:

**No, do not overwrite the existing site. Create a new SharePoint site at** http://<server_name>:port

It is important to note this address because in future this is what you will type into the browser address line to access the site. Press the **Next** button to continue.
The configuration process will now continue. The only option available is to stop the process via the **Cancel** button.

When the process is complete you will see the Window above. Press the **Finish** button.
Chapter 11 – Search

The Search Server Express windows will now automatically be displayed. Click the link **Search Administration** to continue.

The Search Administration site will now be displayed. As you can see it is a WSS v3 site. By default, Search Server Express 2008 will automatically crawl (or index) information in the WSS v3 site, however it is possible to add additional sources for Search Server 2008...
to index. To have additional sources index by Search Server 2008 click on the link **Content Sources** under the **Crawling** item from the left Quick link bar.

You should already see the **Local Office SharePoint Server sites** configured as content sources. Click the **New Content Source** from the menu bar.

Give the new content source a meaningful name, select what sort of your content you wish added. You can choose from:

- SharePoint sites
- Web Sites
- File Shares
- Exchange Public folders

Enter the starting address for the content next. It from this point down that Search Server Express 2008 will index. In this case we have selected a network shared folder.

Scroll down and select the crawl settings, schedules and whether you wish a full crawl to commence immediately. When complete press the OK button to continue.
You should now see the content source you just configured displayed in the list. Continue to add content sources as required following the above method. When complete select the option **Search Administration** from the Quick link bar on the left of the screen under the **Administration** heading.

At the Search Administration page you will now see a summary of all the content sources you have configured and their crawl progress and status.
To commence a search of your configured content simply enter the address of your Search Server site into a browser. Enter a query into the box and click on the magnifying glass or press Enter.

The results from your content will be returned on the left hand side while content matching the search from www.live.com will be returned on the right.

11.9 Indexing .MSG files

Another one of the file types that WSS v3 doesn’t index by default is the .MSG format, which are email messages in Outlook (not Outlook Express) format. This indexing can easily be enabled by the following procedure.
On the WSS v3 server login as an administrator and run the registry editor *regedit*.

Locate the key:

```
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\OfficeServer\12.0\Search\Setup\Filters\.msg
```

You will probably find that it does not exist and therefore needs to be created. To create this key simply **right mouse click** on the *Filters* folder and select **New** then **Key** from the menus that appears.

When the new key has been created name it *.msg* (ensuring that you include the leading period (.)).
Now **right mouse click** on this `.msg` folder and select **New** and **String Value** from the menus that appear.

Enter the value **Extension** as the name of the entry. If you now double click on the entry set the name to be **msg** (no leading period here) and select **OK** to save.
Right mouse click on the .msg key folder again and select New then DWORD value from the menus that appear.

Name this variable FileTypeBucket and set its value to 1 by again double clicking on the entry.

Right mouse click on the .msg key folder once more and select New then String Value from the menus that appear.

The name of this entry is MimeTypes and the value should be application/vnd.ms-outlook (if that doesn’t work you can also try application/msoutlook instead but application/vnd.ms-outlook should work).
The full entry for the new .msg key should appear like shown above. In summary:

Extension = msg
FileTypeBucket = dword:00000001
MimeTypes = application/msoutlook

To complete the process you will need to perform an IISRESET on the WSS v3 server. If you already have .msg files in your WSS v3 site you will also need to perform a full crawl. Details about performing this can be found in Section 11.7 Manual Full Crawl.

You should now find that WSS v3 will index all the content in the Outlook MSG files as all the contents in any attachments associated with these MSG files. Now when you do a search any matching MSG results will be displayed in the search results.

11.10 Conclusion

This guide continues to be a work in progress and I encourage comments and feedback of any type. The only way that the Guide will improve if these is continued feedback.

Please send your comments and feedback to director@ciaops.com.